

**REVIEW****Open Access**

# Major prospects for exploring canine vector borne diseases and novel intervention methods using 'omic technologies

Robin B Gasser<sup>1</sup>\*, Cinzia Cantacessi<sup>1</sup>, Bronwyn E Campbell<sup>1</sup>, Andreas Hofmann<sup>2</sup> and Domenico Otranto<sup>3</sup>

**Abstract**

Canine vector-borne diseases (CVBDs) are of major socioeconomic importance worldwide. Although many studies have provided insights into CVBDs, there has been limited exploration of fundamental molecular aspects of most pathogens, their vectors, pathogen-host relationships and disease and drug resistance using advanced, 'omic technologies. The aim of the present article is to take a prospective view of the impact that next-generation, 'omics technologies could have, with an emphasis on describing the principles of transcriptomic/genomic sequencing as well as bioinformatic technologies and their implications in both fundamental and applied areas of CVBD research. Tackling key biological questions employing these technologies will provide a 'systems biology' context and could lead to radically new intervention and management strategies against CVBDs.

**Background**

Although difficult to estimate, the impact of canine vector-borne diseases (CVBDs) to dog and human populations is substantial [1-3]. Particularly bacteria (e.g., species of *Rickettsia*, *Ehrlichia* and *Borrelia*), protists (including species of *Babesia*, *Leishmania* and *Trypanosoma*), nematodes (e.g., species of *Dirofilaria* and *Acanthocheilonema*), and their vectors (including mosquitoes, fleas, ticks and/or sand flies) constitute major components of the burden of CVBDs [2-4]. With a changing global climate, in the absence of effective preventative approaches and new intervention strategies, the disease burden linked to many neglected CVBDs is likely to increase further [5]. In spite of advances made, there are still major knowledge gaps in CVBDs. These gaps exist mainly in the fundamental molecular biology, epidemiology, ecology and population genetics of causative agents and their vectors, emerging drug resistance issues as well as infection processes and virulence factors [1]. Moreover, substantial limitations in diagnosis and intervention also represent critical obstacles to the effective control of CVBDs. Although sustained research and funding have contributed significantly to an

improved understanding of human vector-borne diseases, such as malaria and trypanosomiasis, this is not the case for many parasitic diseases, which are neglected in terms of research and development [6]. A fundamental change is needed, particularly in relation to CVBDs.

The 'omics era has brought about substantial prospects for investigating some important pathogens and their vectors, providing insights into their epidemiology, ecology, evolution and cellular processes. Available genomes are considered to represent crucial infrastructure for elucidating novel avenues to tackle infectious diseases. However, the relatively high cost and laborious nature of molecular and biochemical research has sometimes been an impediment to progress. Revolutionary developments in a range of 'omic (e.g., genomic, proteomic, metabolomic, glycomic and lipidomic) technologies [7] now provide unprecedented opportunities to explore CVBDs on a scale and at a rate that was unimaginable just a couple of years ago, providing major opportunities for addressing critically important areas of research for the first time ever. Future research should harness such technologies to address major knowledge gaps for CVBDs. Elucidating the pathogens, their relationship with their vector(s) and definitive hosts, the disease(s) as well as the epidemiology and ecology of pathogens causing CVBDs will have substantial prospects to improve the treatment, prevention and control of these parasites

\* Correspondence: robinbg@unimelb.edu.au

<sup>1</sup>Department of Veterinary Science, The University of Melbourne, 250 Princes Highway, Werribee, Victoria 3030, Australia

Full list of author information is available at the end of the article

in years to come. The intent of this presentation is not to review the literature on CVBDs, rather to take a prospective view of the impact that 'omics technologies could have on CVBD research. The emphasis has been placed on describing the principles of transcriptomic/genomic sequencing as well as bioinformatic technologies and their implications in both fundamental and applied areas.

Transcriptomics, for example, is the molecular science of examining, simultaneously, the transcription of all genes at the level of the cell, tissue and/or whole organism, allowing inferences regarding cellular functions and mechanisms. The ability to measure the transcription of thousands of genes simultaneously has led to advances in all biomedical fields, from understanding the basic function in model organisms, such as the yeast, *Saccharomyces cerevisiae* and the vinegar fly, *Drosophila melanogaster* [8-10], to studying molecular processes or mechanisms associated with growth, development and reproduction, to the exploration of the mechanisms of survival and drug-resistance. For more than a decade, transcriptomes have been determined by sequencing expressed sequence tags (ESTs), mainly using a conventional (Sanger) approach [11,12], whereas levels of transcription have been established quantitatively or semi-quantitatively by real-time PCR [13] and/or cDNA microarrays [14]. The use of such technologies has been accompanied by an increasing demand for analytical computer tools for the efficient annotation of nucleotide sequence datasets, particularly within the framework of large-scale EST projects [15]. With a substantial expansion of nucleic acid sequencing has come the development of algorithms for sequence assembly, analysis and annotation, in the form of individual programs [16-18] and integrated pipelines [19,20], some of which have been accessible *via* the worldwide web [19,21,22]. However, the cost and time associated with conventional sequencing and/or the design of customized analytical tools have driven the search for alternative and improved methods [23].

### Next-generation sequencing technologies

There has been an explosion in next-generation sequencing (NGS) technologies [24-27], which now provide unprecedented opportunities to explore *de novo* the transcriptomes and genomes of different species and developmental stages of pathogens, their vectors and/or their definitive hosts. Although introduced recently, the capacity of such techniques to generate millions to hundreds of millions of sequences in parallel has placed them at the forefront of the molecular research [28-30]. Currently available NGS sequencing platforms include 454/Roche [24], Illumina/HiSeq [25] and SOLiD [26].

The 454/Roche platform [24] employs a sequencing-by-synthesis approach, by which cDNA is randomly fragmented (by 'nebulization') into 500-1000 base pair (bp) fragments. For the preparation of the library, an adaptor is ligated to each end of these fragments, which are then mixed with a population of agarose beads whose surfaces anchor oligonucleotides complementary to the 454-specific adapter sequence, such that each bead is associated with a single fragment. Each of these complexes is transferred into individual oil-water micelles containing amplification reagents and is then subjected to an emulsion PCR (emPCR) step, during which ~10 million copies of each cDNA are produced and bound to individual beads. In the sequencing phase, the beads anchoring the cDNAs are deposited on a pico-titre plate, together with other enzymes required for the pyrophosphate sequencing reaction (i.e., ATP sulfurylase and luciferase). The sequencing is carried out by flowing sequencing reagents (nucleotide and buffers) over a plate [31]. To date, the 454 sequencing technology is a 'long-read' (100-600 bp) platform and is often used for *de novo* genomic or transcriptomic studies.

The Illumina/HiSeq (formerly Solexa) technology has features that differ significantly from the 454 approach [25]. After fragmentation of cDNA sample into a shotgun library, Illumina-specific adaptors are ligated *in vitro* to each cDNA template; one terminus of the template is covalently attached to the surface of a glass slide (or flow cell). Attached to the flow cell are primers complementary to the other end of the template, which bend the cDNAs to form bridge-like structures. During the amplification step (bridge-PCR), clonal clusters, each consisting of ~1000 amplicons, are generated and immobilized to a single physical location on the slide. Subsequently, the cDNAs are linearised, and the sequencing reagents are directly added to the flow cell, with four fluorescently labelled nucleotides. After the incorporation of individual fluorescent bases, the flow cell is interrogated with a laser in several locations, which results in several image acquisitions at the end of a single synthesis cycle [31]. This technology is considered ideal for re-sequencing projects, targeted sequencing, single nucleotide polymorphism (SNP) analyses and gene transcription studies.

The SOLiD platform employs the enzyme DNA ligase, instead of a polymerase [26]. Briefly, after an emPCR step, the adaptor sequences of the cDNA templates bind to complementary primers that are covalently anchored to a glass slide. Subsequently, a set of four fluorescently labelled di-probes (octamers of random sequence, except known dinucleotides at the 3'-terminus) is added to the sequencing reaction. In case an octamer is complementary to the template, it is ligated, and the two specific nucleotides can be called;

subsequently, an image is acquired and the fluorescent dye is removed, so that other octamers can be ligated. After multiple ligations (e.g., 7 ligations for a 35 bp read), the newly synthesized cDNA is removed and the primer is inactivated. This process is repeated multiple times from different starting points of the cDNA templates, so that each position is sequenced at least twice ('two-base-calling'). Because of the short read-length, the range of applications of the SOLiD system is considered similar to that of the Illumina technology and includes (targeted) re-sequencing projects, SNP detection and gene transcription studies.

In the last years, a range of studies have demonstrated the utility of NGS technologies for investigating, for instance, aspects of the molecular biology, systematics and population genetics of parasites [32-36]. In particular, 454 technology was used recently for the rapid *de novo* sequencing of the transcriptomes of numerous pathogens of humans and animals [33-41], yielding substantial datasets and providing a significant step forward. The development of practical and efficient bioinformatic tools has now become crucial for comprehensive analyses of such datasets.

### Bioinformatic tools for the analysis of sequence datasets

The application of NGS technologies has been accompanied by an expansion of bioinformatic tools for the analysis of DNA, RNA and protein sequence datasets. This expansion has resulted in the development of a number of programs and/or integrated pipelines accessible via the world-wide web (www) (e.g., [16,19,21,42,43]). The principles, methods and protocols for the analysis of sequence data, together with currently available bioinformatic tools and pipelines, have been reviewed [42].

In brief, following the acquisition of data, sequences are firstly screened for repeats, contaminants and/or adaptor sequences [42,44] and 'clustered' (= assembled) into contiguous sequences (of maximum length; called contigs) based on sequence similarity (see [42]). Long (e.g., generated by Sanger sequencing, 454 platform) and short-reads (e.g., Illumina and SOLiD platforms) are assembled using the algorithms 'overlap-layout-consensus' [45] and 'de Bruijn graph' [46,47], respectively. For the former algorithm [45], pair-wise overlaps among reads are computed and stored in a graph; all graphs are then used to compute a layout of reads and a consensus sequence of contigs [21,48-53]. For the 'de Bruijn graph' [46,47], reads are fragmented into short segments, denominated 'k-mers', where 'k' represents the number of nucleotides in each segment; overlaps between or among k-mers are captured and stored in graphs, which are subsequently used to generate the consensus sequences [47,52-56].

Following assembly, the contigs and single reads (or singletons) are compared, using different types of the Basic Local Alignment Software Tool (BLAST; [57]) with known sequence data available in public databases, in order to assign a predicted identity to each query sequence if significant matches are found [42]. In addition, assembled nucleotide sequences are usually conceptually translated into predicted proteins using algorithms that identify protein-coding regions (ORFs) from individual contigs [22,58,59]. Once peptides are predicted, protein analyses, including amino acid sequence comparisons with data available in public databases, and known protein domains, are then undertaken [17,42,60-63]. Public databases (accessible via www) represent comprehensive collections of nucleotide and amino acid sequences, where all publicly available nucleotide sequences are stored and curated [64-66]; in addition, each sequence is stored as a separate record and linked to salient information, such as primary source, references and predicted and/or experimentally verified biological features. For transcriptomic datasets, examples of databases include the UniGene [67] and the Sequence Read Archive (SAR) [68]. In addition to these general databases, there are various specialized collections of gene and protein information on particular (model) organisms about which much is known. Examples include the databases for *S. cerevisiae* (yeast; <http://www.yeastgenome.org/>) [69], *D. melanogaster* (vinegar fly; <http://flybase.org/>) [70], *Mus musculus* (mouse; <http://www.informatics.jax.org/>) [71] and *Caenorhabditis elegans* (free-living nematode; WormBase at <http://www.wormbase.org/>) [72,73].

A web-based bioinformatic pipeline (= ESTExplorer) was established for the automated analysis and annotation of nucleic acid datasets (both at the nucleotide and amino acid levels) [19], and shown to substantially accelerate and facilitate the analyses of sequences (generated using conventional Sanger sequencing) compared with traditional database searches [20]. However, sequences generated by NGS are significantly shorter (454/Roche: ~400 bases; Illumina/SOLiD: ~60 bases) than those determined by Sanger sequencing (0.8-1 kb), which poses a significant challenge for assembly. In addition, the data files generated by these technologies are often gigabytes to terabytes ( $1 \times 10^9$  to  $1 \times 10^{12}$  bytes) in size, substantially increasing the demands placed on data transfer and storage, such that most web-based interfaces are no longer suited for large-scale analyses. In order to overcome this limitation, a recent report [39] described the development of an integrated bioinformatic workflow system for the analysis and annotation of large sequence datasets produced by NGS, in which the majority of the software was derived from existing application tools (e.g., CAP3; [21]), available as

web-based interfaces. These applications, optimized using the Linux operation system, were incorporated into pre-existing scripts (i.e., Perl, Python and Unix shell), and can be downloaded <http://research.vet.unimelb.edu.au/gasserlab/index.html>[39] and readily executed, also by scientists with limited bioinformatic expertise. This workflow system has proved very useful for time-efficient and accurate analyses of large-scale transcriptomic datasets generated by NGS and for distilling biologically meaningful information (such as predictions of essential molecules) on the parasite, the vector or the host under investigation.

### **Exciting prospects in both fundamental and applied areas**

Knowledge of the transcriptomes and proteomes of different developmental stages of a parasite, its vector and its definitive host is central to gaining an enhanced understanding of the molecular mechanisms that govern essential biological, infection and disease processes and, ultimately, could assist in identifying possible avenues for the development of novel intervention strategies. Accurate bioinformatic analyses of nucleic acid and protein sequence data (often by comparison with or inference from reference organisms) are crucial, in the absence of information for the organism under study, in providing biological meaningful molecular biological information about CVBDs. Until recently, detailed bioinformatic analyses of such datasets have been restricted largely to specialized laboratories with substantial computer and software capacities. The development of flexible and practical bioinformatic workflow systems is beginning to provide scientists with user-friendly tools for the analysis of massive datasets.

Currently, due to a lack of complete genomic sequences for many pathogens and vectors (and different strains thereof) associated with CVBDs, newly generated sequence datasets need to be assembled *de novo*, which means that pooled reads are assembled without a bias towards known sequences [47]. Due to the amount of RNA required for NGS (~5-10 µg) [74], transcriptomes usually originate from numerous individuals, potentially leading to an increased complexity of the sequence data acquired (linked, for instance, to a biased nucleotide content, single nucleotide polymorphisms [SNPs] and other types of sequence variation) and sometimes posing challenges for the data assembly. In terms of complexity, computational and time requirements, *de novo* assemblies are much slower and more computer-memory intensive than knowledge-based (mapping) assemblies, in which reads are aligned and assembled against an existing reference sequence (representing the same species or genetic variant) [18]. In addition, reliable *de novo* assemblies are highly

dependent upon the availability of long reads (>100 bases) and of high-coverage, paired-end sequence data [75]. In previous studies, the complementary nature of the 454 and Illumina sequencing platforms has allowed the assembly of raw reads into large scaffolds without a need for a reference sequence [76-78].

In the absence of reference genomes for agents and vectors linked to CVBDs, accurate assembly of sequence data is a crucial step in examining coding genes and, ultimately, addressing biological questions regarding gene and protein functions. Functions are initially predicted by 'sequence annotation' (= the process of gathering all available information and relating it to the sequence assembly both by experimental and computational means [79]. Accurate annotation is dependent on the efficiency of the updates and curation. Presently, open-source programs and databases routinely employed for the bioinformatic analyses of sequence data are available *via* multiple portals, thus requiring significant efforts to maintain accurate and up-to-date assembly and annotation pipelines [80]. In addition, the rate at which public databases are updated and corrected varies considerably. For instance, the Swiss-Prot database <http://au.expasy.org/sprot/> accepts corrections from its user community, whereas GenBank <http://www.ncbi.nlm.nih.gov/genbank/> only accepts corrections from the author of an entry [81], thus significantly affecting the accuracy and speed with which new sequences are annotated. In addition, some information-management systems incorporate data from large-scale projects, but often, the annotation of single records from the literature is slow [82]. Given that, presently, the annotation of sequence data for parasites and vectors relies heavily on the use of bioinformatic approaches and already annotated/curated sequence data for a wide range of organisms, these aspects deserve careful consideration.

The analyses and annotation of large-scale transcriptomic, proteomic and genomic sequence datasets for pathogens could be facilitated through the establishment of a 'reference' website for CVBDs. Such a website could provide regular releases of newly developed and validated bioinformatic pipelines for the analyses of sequence datasets. It could also provide links to regularly updated databases that are routinely employed for the annotation of new sequences as well as a distinct, high-quality database of curated functional annotations, supported by experimental data published in peer-reviewed, international publications. In the future, the establishment of a 'centralized' resource to enable the sharing and optimization of bioinformatic pipelines for sequence processing and annotation and, more broadly, to allow access to new sequence data, and experimental protocols and relevant literature would be advantageous.

The annotation of peptides inferred from a dataset is conducted by assigning predicted biological function/s based on comparison with existing information available for related organisms in public databases, including InterPro <http://www.ebi.ac.uk/interpro/>, Gene Ontology, <http://www.geneontology.org/>, OrthoMCL <http://www.orthomcl.org/>, BRENDA <http://www.brenda-enzymes.org/>. Using this approach, predictions for key groups of molecules can be made regarding their fundamental functional and essential roles in biological processes [61]. Such groups include molecules linked to the physiology of the nervous system [37], the formation of the cuticle (arthropods and nematodes) [37,83], reproduction, development, signal transduction and/or pathogen invasion and disease processes (e.g., proteases and protease inhibitors, protein kinases and phosphatases) [36-39,84].

The bioinformatic prediction and prioritization of novel drug targets involves 'filtering' [85,86] and usually includes inferring targets based on key principles and requirements [87-91]. First, target proteins should have one or more essential roles in fundamental biological processes of the pathogen and/or vector, such that the disruption of the molecule or its gene will damage and/or kill both or either and thus disrupt disease transmission or disease itself, but not affect the host [90,92]. In the absence of phenotypic data for many pathogens/vectors, the prediction of drug target candidates in eukaryotic pathogens/vectors can be assisted by using extensive information on function and essentiality in a range of eukaryotic organisms, including *S. cerevisiae*, *D. melanogaster*, *C. elegans* and *M. musculus*. This information can be accessed via public databases, including FlyBase at <http://flybase.org/>, WormBase at <http://www.wormbase.org>, Mouse Genome Informatics at <http://www.informatics.jax.org/> and *Saccharomyces* Genome Database at <http://www.yeastgenome.org/> [39,89,93-95]. Since most effective drugs achieve their activity by competing with endogenous small molecules for a binding site on a target protein [96], the amino acid sequences predicted from essential genes should be screened for the presence of relatively conserved ligand-binding domains [96,97]. Lists of inhibitors, known based on experimental evidence, to specifically bind to such domains, can be compiled. However, the predictions made are intended to support hypothesis-driven or applied research and thus require extensive experimental investigations. The main advantage for a number of CVBD pathogens (e.g., *Babesia* and *Leishmania*) over, for example, some parasitic helminths, is that they can be propagated readily *in vitro* (e.g., [98,99]). This provides unique prospects to test gene function(s) by double-stranded RNA interference, transgenesis and/or deletion studies as well as using small molecular inhibitors (cf. [100-102]).

Based on recent evidence [103-105], guanosine triphosphatases (GTPases), protein phosphatases and protein kinases seem to represent attractive drug target candidates for a range of pathogens, but have not yet been examined on a genome-wide scale and in a systematic manner for most CVBDs. Multiple cellular signaling pathways function through the activity of small GTP-binding proteins to regulate multiple biological processes, such as transmembrane signal transduction, cytoskeletal reorganization, gene expression, intracellular vesicle trafficking, microtubule organization and nucleocytoplasmic transport [106]. GTPases are small (~20-28 kDa), monomeric proteins belonging to six families (i.e., Ras, Rho, Rab, Arf, Ran and Rad; [107]). These regulatory proteins act as bi-molecular switches that cycle between two conformational states (i.e., GDP-bound ['inactive' state] and GTP-bound ['active' state]) and hydrolyze GTP. In humans, the aberrant regulation of GTPases is linked to a number of dysfunctions, including neurological and developmental disorders and cancer [108]. In addition, intracellular pathogenic bacteria, such as *Mycobacterium tuberculosis*, are known to target host GTPases to evade host immune responses to facilitate the infection process [109]. Such information has stimulated efforts to develop novel therapeutic strategies to inhibit the function of GTPases. For instance, treatments with farnesyltransferase inhibitors, to block the oncogenic properties of Ras GTPases, have been shown to be effective in significantly reducing the progression of various forms of cancer, including carcinomas of the colon, pancreas and lung, neurofibrosarcoma and chronic myelogenous leukaemia, in experimental animals [110,111] and the migration and organization of the cytoskeleton of human prostate cancer cells [112]. Although the overall structure of individual small GTPases is conserved across eukaryotes, the filtering of datasets for the organism of interest (i.e., pathogen and/or vector) allows the identification of significant differences in sequence of GTPases between the invertebrate and the definitive host. These differences might be considered in future studies, aimed at assessing the possibility of designing and synthesizing selective and specific inhibitors against parasite GTPases. Homology modelling [113,114], X-ray crystallography/nuclear magnetic resonance (NMR) and docking [115-120] studies should assist in this process.

Selected protein kinases (PKs) are also potential drug targets for a range of pathogens. PKs belong to a large family of proteins regulating development, cell division, differentiation and metabolism in many organisms; these molecules are considered the second most important group of drug targets after GPCRs [121,122]. The family of PKs comprises cell surface receptors and non-receptor or cytosolic kinases. Integrated genomic-bioinformatic-chemoinformatic approaches have been employed

for the identification and screening effective PK inhibitors as therapeutic agents [123-125]. For example, in studies aimed at identifying novel inhibitors of a human tyrosine kinase involved in the development and progression of chronic myelogenous leukemia, 15 compounds were selected following *in silico* screening of a database of 200,000 known inhibitors [124]. Of these compounds, eight were shown to selectively inhibit the growth of leukemia *in vitro* [124]. In another study, novel and selective inhibitors of caseine kinase II (CK2) were identified via *in silico* screening of a database containing ~400,000 compounds, followed by *in silico* docking [125]. These examples indicate the advantages of using computer-aided tools for the rational prediction and design of drugs for subsequent *in vitro* and *in vivo* efficacy testing [126]. Nonetheless, it is clear that any compound shown to be efficacious must also be rigorously tested for its safety (see [127]; <http://www.ich.org/cache/compo/276-254-1.html>).

Because of the regulatory role that PKs play in a number of signaling pathways in the cell, interference with their activity can result in the disruption of fundamental homeostatic processes in parasites [105]. In the last years, protein kinases have received particular attention as drug targets in protists, such as species of *Plasmodium*, *Leishmania* and *Trypanosoma* and helminths [105]. For instance, particular inhibitors of pyrrole and imidazopyridine cyclic guanosine monophosphate-dependent protein kinases of *Leishmania major* have been shown to severely impair the growth of the promastigote forms of this protozoan parasite *in vitro* [128]. In some helminths, for example, PK inhibitors (i.e., tyrophostins AG1024 and AG538) have been shown to significantly affect the survival and development of the adult parasite through the blockage of glucose uptake [122]. The inactivation of PKs with herbimycin A has also been shown to interfere with mitosis, thus significantly affecting the expression of proteins essential for egg production in the worm [129]. Although the crystal structures of PKs in many pathogens have not yet been defined, progress has been made in the identification and design of effective inhibitors based on homology models for protein kinases from humans [105]. There is evidence that the active sites of parasite PKs display subtle differences compared with their human counterparts [105], which is considered promising for the development of parasite-specific kinase inhibitors. However, much more study is required to establish the potential of PK inhibitors against pathogens causing CVBDs. This is obviously a research area worth pursuing.

### Concluding remarks

Vector-borne diseases, of which CVBDs represent a substantial component, represent ~17% of the burden of all

infectious diseases and have a major socioeconomic impact worldwide [130]. In addition to their veterinary importance, some CVBD-causing agents are of major zoonotic importance. Although various studies have provided improved insights into the epidemiology of CVBDs using molecular methods, there has been limited study of fundamental molecular aspects of many pathogens, their vectors, pathogen-host relationships and disease as well as drug or insecticide resistance using some of the advanced 'omic technologies described here. Tackling fundamental biological questions using these technologies and understanding the relationship among pathogens/vectors/environment will provide a 'systems biological context' to better understand CVBDs and their epidemiology and should lead to the design of radically new intervention and management strategies against these diseases.

For instance, from a fundamental perspective, genomic sequencing and the definition of a wide range of genetic markers for use in specific and sensitive diagnostic tools could provide a solid foundation for addressing questions regarding the complex network of biological and ecological factors involved in pathogen/host/environment interactions and the immunological idiosyncrasies of receptive hosts in endemic regions as well as the role of asymptomatic, chronically infected animals and those infected with multiple pathogens [2]. In this context, using well-defined genetic and transcriptomic tools, it would be interesting to address the question as to whether simultaneous infections with multiple vector-borne pathogens (compared with a single infection) induce synergistic and pronounced immunosuppression in infected animals. Moreover, the application of -omics tools could also assist in comprehensively studying the complex intracellular pathways that are manipulated or regulated by one or multiple pathogens (e.g., species of *Leishmania* and *Ehrlichia*) to evade the immune response of the host and significantly complicating the progression and expression of disease in individual patients (cf. [131,132]). It would also be very useful to investigate the resistance and susceptibility of, for example, particular dog breeds to CVBD-agents and their vectors. For instance, a genomic comparison between Ibizan hounds (which are resistant to leishmaniasis; [133]) and other breeds, such as Boxers, as well as transcriptomic/proteomic comparisons of the responses of these dogs to infection and disease would be very interesting. In a broader context, gaining improved insights into the relationship between host genotype (through genomic sequencing) and phenotype (degree of disease expression) in response to particular CVBD-pathogens and/or intervention approach (e.g., treatment/vaccination) would be particularly informative and could assist in a deeper understanding the genetic basis of disease.

From an epidemiological perspective, also changes in the spatial and temporal distribution of pathogens, vectors and/or their hosts, as a result of climatic change and global warming, might also be monitored using metagenomic approaches. These examples indicate clearly that there are many exciting fundamental areas to tackle using genomic, proteomic and immunomic tools in the very near future.

From an applied perspective, clearly, the improved prediction and prioritization of drug and vaccine targets in CVBD pathogens or repellants against vectors is a key area. NGS will provide the efficiency and depth-of-coverage required to rapidly define *de novo* the complete genomes of hosts, CVBD pathogens and their vectors. Repertoires of drug or vaccine targets could be inferred on a global scale. For example, the parasite kinome (= the complete set of kinase genes in the genome) could represent a unique opportunity for the design of pathogen-selective inhibitors [105] for subsequent validation by high throughput screening of parasites [134–137]. The combined use of genomic, transcriptomic, proteomic and metabolomic datasets will be crucial to identifying other groups of molecules essential to the development and survival of a pathogen for the design of novel classes of small molecular inhibitors. Clearly, an integrated use of 'omic technologies will pave the way to a better understanding of the systems biology of CVBDs on a scale never before possible, and, hopefully, will provide golden opportunities for the development of entirely new intervention strategies in public-private partnerships.

#### Acknowledgements

Bayer Animal Health GmbH sponsored the thematic series for the Symposium CVBD6. RBG was an invited, plenary speaker. Current research in the Gasser Lab is funded mainly by the Australian Research Council (ARC), the National Health & Medical Research Council (NHMRC) and Melbourne Water Corporation.

#### Author details

<sup>1</sup>Department of Veterinary Science, The University of Melbourne, 250 Princes Highway, Werribee, Victoria 3030, Australia. <sup>2</sup>Structural Chemistry Program, Eskitis Institute for Cell & Molecular Therapies, Griffith University, Brisbane, Queensland, Australia. <sup>3</sup>Dipartimento di Sanità Pubblica e Zootecnia, Facoltà di Medicina Veterinaria, Università di Bari, Str. prov. le per Casamassima Km 3, 70010, Valenzano, Bari, Italy.

#### Authors' contributions

RBG, CC, BEC, AH and DO all contributed actively to the drafting of the manuscript.

#### Competing interests

The authors declare that they have no competing interests.

Received: 24 January 2011 Accepted: 13 April 2011

Published: 13 April 2011

#### References

- Otranto D, Wall R: New strategies for the control of arthropod vectors of disease in dogs and cats. *Med Vet Entomol* 2008, **22**:291–302.
- Otranto D, Dantas-Torres F, Breitschwerdt EB: Managing canine vector-borne diseases of zoonotic concern: part one. *Trends Parasitol* 2009, **25**:157–163.
- Otranto D, Dantas-Torres F, Breitschwerdt EB: Managing canine vector-borne diseases of zoonotic concern: part two. *Trends Parasitol* 2009, **25**:228–235.
- Otranto D, Dantas-Torres F: Canine and feline vector-borne diseases in Italy: current situation and perspectives. *Parasit Vectors* 2010, **3**:2.
- Mills JN, Gage KL, Khan AS: Potential influence of climate change on vector-borne and zoonotic diseases: A review and proposed research plan. *Environ Health Perspect* 2010.
- Moran M, Guzman J, Ropars AL, McDonald A, Jameson N, Omune B, Ryan S, Wu L: Neglected disease research and development: how much are we really spending? *PLoS Med* 2009, **6**:e30.
- Walker JM, Rapley R: *Molecular Biology and Biotechnology*. Cambridge: The Royal Society of Chemistry Press; 2009.
- Stathopoulos A, Levine M: Whole-genome expression profiles identify gene batteries in *Drosophila*. *Dev Cell* 2002, **3**:464–465.
- Gupta V, Oliver B: *Drosophila* microarray platforms. *Brief Funct Genomic Proteomic* 2003, **2**:97–105.
- Vibranski MD, Lopes HF, Karr TL, Long M: Stage-specific expression profiling of *Drosophila* spermatogenesis suggests that meiotic sex chromosome inactivation drives genomic relocation of testis-expressed genes. *PLoS Genet* 2009, **5**:e1000731.
- Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 1977, **74**:5463–5467.
- Sanger F, Air GM, Barrell BG, Brown NL, Coulson AR, Fiddes CA, Hutchinson CA, Slocombe PM, Smith M: Nucleotide sequence of bacteriophage phi X174 DNA. *Nature* 1977, **265**:687–695.
- Wang AM, Doyle MV, Mark DF: Quantitation of mRNA by the polymerase chain reaction. *Proc Natl Acad Sci USA* 1989, **86**:9717–9721.
- DeRisi J, Penland L, Brown PO, Bittner ML, Meltzer PS, Ray M, Chen Y, Su YA, Trent JM: Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat Genet* 1996, **14**:457–460.
- Clifton SW, Mitreva M: Strategies for undertaking expressed sequence tag (EST) projects. *Methods Mol Biol* 2009, **533**:13–32.
- Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M: Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 2005, **21**:3674–3676.
- Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Das U, Daugherty L, Duquenne L, Finn RD, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Laugraud A, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, Mistry J, Mitchell A, Mulder N, Natale D, Orengo C, Quinn AF, Selengut JD, Sigrist CJ, Thimma M, Thomas PD, Valentin F, Wilson D, Wu CH, Yeats C: InterPro: the integrative protein signature database. *Nucleic Acids Res* 2009, **37**:D211–D215.
- Flicek P, Birney E: Sense from sequence reads: methods for alignment and assembly. *Nat Methods* 2009, **6**:S6–S12.
- Nagaraj SH, Deshpande N, Gasser RB, Ranganathan S: ESTExplorer: an expressed sequence tag (EST) assembly and annotation platform. *Nucleic Acids Res* 2007, **35**:W143–W147.
- Nagaraj SH, Gasser RB, Nisbet AJ, Ranganathan S: In silico analysis of expressed sequence tags from *Trichostrongylus vitrinus* (Nematoda): comparison of the automated ESTExplorer workflow platform with conventional database searches. *BMC Bioinf* 2008, **9**:510.
- Huang X, Madan A: CAP3: A DNA sequence assembly program. *Genome Res* 1999, **9**:868–877.
- Iseli C, Jongeneel CV, Bucher P: ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *Proc Int Conf Intell Syst Mol Biol* 1999, **1**:138–148.
- Morozova O, Marra MA: Applications of next-generation sequencing technologies in functional genomics. *Genomics* 2008, **92**:255–264.
- Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irsik GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF,

Rothberg JM: **Genome sequencing in microfabricated high-density picolitre reactors.** *Nature* 2005, **437**:376-380.

25. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Keira Cheetham R, Cox AJ, Ellis DJ, Flatbush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS, Kirk SM, Li H, Liu X, Maisinger KS, Murray LJ, Obradovic B, Ost T, Parkinson ML, Pratt MR, Rasolonjatovo IM, Reed MT, Rigatti R, Rodighiero C, Ross MT, Sabot A, Sankar SV, Scally A, Schroth GP, Smith ME, Smith VP, Spiridou A, Torrance PE, Tzanev SS, Vermaas EH, Walter K, Wu X, Zhang L, Alam MD, Anastasi C, Aniebo IC, Bailey DM, Bancarz IR, Banerjee S, Barbour SG, Baybayan PA, Benoit VA, Benson KF, Bevis C, Black PJ, Bodhun A, Brennan JS, Bridgman JA, Brown RC, Brown AA, Buermann DH, Bundu AA, Burrows JC, Carter NP, Castillo N, Chiara E, Catenazzi M, Chang S, Neil Cooley R, Crane NR, Dada OO, Diakoumaki KD, Dominguez-Fernandez B, Earnshaw DJ, Egbojor UC, Elmore DW, Etchin SS, Ewan MR, Fedurco M, Fraser LJ, Fuentes Fajardo KV, Scott Furey W, George D, Gietzen KJ, Goddard CP, Golda GS, Granieri PA, Green DE, Gustafson DL, Hansen NF, Harnish K, Haudenschild CD, Heyer NI, Hims MM, Ho JT, Horgan AM, Hoschler K, Hurwitz S, Ivanov DV, Johnson MQ, James T, Huw Jones TA, Kang GD, Kerelska TH, Kersey AD, Kherbukova I, Kindwall AP, Kingsbury Z, Kokko-Gonzales PI, Kumar A, Laurent MA, Lawley CT, Lee SE, Lee X, Liao AK, Loch J, Lok M, Luo S, Mammen RM, Martin JW, McCauley PG, McNitt P, Mehta P, Moon KW, Mullens JW, Newtonton T, Ning Z, Ling Ng B, Novo SM, O'Neill MJ, Osborne MA, Osnowski A, Ostdan O, Paraschos LL, Pickering L, Pike AC, Pike AC, Chris Pinkard D, Pliskin DP, Podhasky J, Quijano VJ, Raczy C, Rae VH, Rawlings SR, Chiva Rodriguez A, Roe PM, Rogers J, Robert Bacigalupo MC, Romanov N, Romieu A, Roth RK, Rourke NJ, Ruediger ST, Rusman E, Sanchez-Kuiper RM, Schenker MR, Seoane JM, Shaw RJ, Shiver MK, Short SW, Sizto NL, Sluis JP, Smith MA, Ernest Sohna Sohna J, Spence EJ, Stevens K, Sutton N, Szajkowski L, Tregidgo CL, Turcatti G, Vandevondele S, Verhovsky Y, Virk SM, Wakelin S, Walcott GC, Wang J, Worsley GJ, Yan J, Yau L, Zuerlein M, Rogers J, Mullikin JC, Hurles ME, McCooke NJ, West JS, Oaks FL, Lundberg PL, Klenerman D, Durbin R, Smith AJ: **Accurate whole human genome sequencing using reversible terminator chemistry.** *Nature* 2008, **456**:53-59.

26. Harris TD, Buzby PR, Babcock H, Beer E, Bowers J, Braslavsky I, Causey M, Colonell J, Dimeo J, Efcaitch JW, Giladi E, Gill J, Healy J, Jarosz M, Lapan D, Moulton K, Quake SR, Steinmann K, Thayer E, Tyurina A, Ward R, Weiss H, Xie Z: **Single-molecule DNA sequencing of a viral genome.** *Science* 2008, **320**:106-109.

27. Pandey V, Nutter RC, Prediger E: **Applied Biosystems SOLiD™ System: Ligation-Based Sequencing.** In *Next Generation Genome Sequencing: Towards Personalized Medicine*. Edited by: Milton JM. Australia: Wiley; 2008:29-41.

28. Mardis ER: **The impact of next-generation sequencing technology on genetics.** *Trends Genet* 2008, **24**:133-141.

29. Wang Z, Gerstein M, Snyder M: **RNA-seq: a revolutionary tool for transcriptomics.** *Nat Rev Genet* 2009, **10**:57-63.

30. Marguerat S, Bahler J: **RNA-seq: from technology to biology.** *Cell Mol Life Sci* 2010, **67**:569-579.

31. Mardis ER: **Next-generation DNA sequencing methods.** *Ann Rev Genom Human Genet* 2008, **9**:387-402.

32. Jex AR, Hu M, Littlewood DT, Waeschenbach A, Gasser RB: **Using 454 technology for long-PCR based sequencing of the complete mitochondrial genome from single *Haemonchus contortus* (Nematoda).** *BMC Genomics* 2008, **9**:11.

33. Young ND, Campbell BE, Hall RS, Jex AR, Cantacessi C, Laha T, Sohn WM, Sripa B, Loukas A, Brindley PJ, Gasser RB: **Unlocking the transcriptomes of two carcinogenic parasites, *Clonorchis sinensis* and *Opisthorchis viverrini*.** *PLoS Negl Trop Dis* 2010, **4**:e719.

34. Young ND, Hall RS, Jex AR, Cantacessi C, Gasser RB: **Elucidating the transcriptome of *Fasciola hepatica*-a key to fundamental and biotechnological discoveries for a neglected parasite.** *Biotechnol Adv* 2010, **28**:222-231.

35. Young ND, Jex AR, Cantacessi C, Campbell BE, Laha T, Sohn WM, Sripa B, Loukas A, Brindley PJ, Gasser RB: **Progress on the transcriptomics of carcinogenic liver flukes of humans-unique biological and biotechnological prospects.** *Biotechnol Adv* 2010, **28**:859-870.

36. Wang Z, Abubucker S, Martin J, Wilson RK, Hawdon J, Mitreva M: **Characterizing *Ancylostoma caninum* transcriptome and exploring nematode parasitic adaptation.** *BMC Genomics* 2010, **11**:307.

37. Cantacessi C, Campbell BE, Young ND, Jex AR, Hall RS, Presidente PJ, Zawadzki JL, Zhong W, Aleman-Meza B, Loukas A, Sternberg PW, Gasser RB: **Differences in transcription between free-living and CO<sub>2</sub>-activated third-stage larvae of *Haemonchus contortus*.** *BMC Genomics* 2010, **11**:266.

38. Cantacessi C, Mitreva M, Jex AR, Young ND, Campbell BE, Hall RS, Doyle MA, Ralph SA, Rabelo EM, Ranganathan S, Sternberg PW, Loukas A, Gasser RB: **Massively parallel sequencing and analysis of the *Necator americanus* transcriptome.** *PLoS Negl Trop Dis* 2010, **4**:e684.

39. Cantacessi C, Jex AR, Hall RS, Young ND, Campbell BE, Joachim A, Nolan MJ, Abubucker S, Sternberg PW, Ranganathan S, Mitreva M, Gasser RB: **A practical, bioinformatic workflow system for large data sets generated by next-generation sequencing.** *Nucleic Acids Res* 2010, **38**:e171.

40. Cantacessi C, Mitreva M, Campbell BE, Hall RS, Young ND, Jex AR, Ranganathan S, Gasser RB: **First transcriptomic analysis of the economically important parasitic nematode, *Trichostrongylus colubriformis*, using a next-generation sequencing approach.** *Infect Genet Evol* 2010, **10**:1199-1207.

41. Cantacessi C, Gasser RB, Strube C, Schnieder T, Jex AR, Hall RS, Campbell BE, Young ND, Ranganathan S, Sternberg PW, Mitreva M: **Deep insights into *Dictyocaulus viviparus* transcriptomes provides unique prospects for new drug targets and disease intervention.** *Biotechnol Adv* 2011.

42. Nagaraj SH, Gasser RB, Ranganathan S: **A hitchhiker's guide to expressed sequence tag (EST) analysis.** *Brief Bioinformatics* 2007, **8**:6-21.

43. Soderlund C, Johnson E, Bomhoff M, Descour A: **PAVE: program for assembling and viewing ESTs.** *BMC Genomics* 2009, **10**:400.

44. Falqueras J, Lara AJ, Fernandez-Poso N, Canton FR, Perez-Trabado G, Clarov MG: **SeqTrim: a high throughput pipeline for pre-processing any type of sequence read.** *BMC Bioinformatics* 2010, **11**:38.

45. Myers EW: **Toward simplifying and accurately formulating fragment assembly.** *J Comput Biol* 1995, **2**:275-290.

46. Idury RM, Waterman MS: **A new algorithm for DNA sequence assembly.** *J Comput Biol* 1995, **2**:291-306.

47. Zerbino DR, Birney E: **Velvet: algorithms for de novo short read assembly using de Bruijn graphs.** *Genome Res* 2008, **18**:821-829.

48. Sutton GG, White O, Adams MD, Kerlavage AR: **TIGR assembler: a new tool for assembling large shotgun sequencing projects.** *Genome Sci Technol* 1995, **1**:9-19.

49. Green P: **Documentation for PHRAP.** Genome Center, University of Washington, Seattle. 1996 [<http://www.phrap.org/phredphrap/phrap.html>].

50. Huang X, Wang J, Aluru S, Yang SP, Hillier L: **PCAP: a whole genome assembly program.** *Genome Res* 2003, **13**:2164-2170.

51. Chevreux B: **MIRA: an automated genome and EST assembler.** Ph.D Thesis, German Cancer Research Center Heidelberg, Duisburg, Germany. 2005.

52. Scheibye-Alsing K, Hoffmann S, Frankel A, Jensen P, Stadler PF, Mang Y, Tommerup N, Gilchrist MJ, Nygård AB, Cirera S, Jørgensen CB, Fredholm M, Gorodkin J: **Sequence assembly.** *Comput Biol Chem* 2009, **33**:121-136.

53. Miller JR, Koren S, Sutton G: **Assembly algorithms for next-generation sequencing data.** *Genomics* 2010, **95**:315-327.

54. Warren RL, Sutton GC, Jones SJM, Holt RA: **Assembling millions of short DNA sequences using SSAKE.** *Bioinformatics* 2007, **23**:500-501.

55. Hernandez D, François P, Farinelli L, Osteras M, Schrenzel J: **De novo bacterial genome sequencing: millions of very short reads assembled on a desktop computer.** *Genome Res* 2008, **18**:802-809.

56. Li R, Li Y, Kristiansen K, Wang J: **SOAP: short oligonucleotide alignment program.** *Bioinformatics* 2008, **24**:713-714.

57. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic Local Alignment Search Tool.** *J Mol Biol* 1990, **215**:403-410.

58. Fukunishi Y, Hayashizaki Y: **Amino acid translation program for full-length cDNA sequences with frameshift errors.** *Physiol Genomics* 2001, **5**:81-87.

59. Min XJ, Butler G, Storms R, Tsang A: **OrfPredictor: predicting protein-coding regions in EST-derived sequences.** *Nucleic Acids Res* 2005, **33**:W677-W680.

60. Hofmann K, Bucher P, Falquet L, Bairoch A: **The Prosite Database, its status in 1999.** *Nucleic Acids Res* 1999, **27**:215-219.

61. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G: **Gene Ontology: tool for the unification of biology.** *The Gene Ontology Consortium.* *Nat Genet* 2000, **25**:25-29.

62. Attwood TK, Croning MD, Flower DR, Lewis AP, Mabey JE, Scordis P, Selley JN, Wright W: **Prints-S: the database formerly known as prints.** *Nucleic Acids Res* 2000, **28**:225-227.

63. Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, Sonnhammer EL: **The Pfam protein families database.** *Nucleic Acids Res* 2000, **28**:263-266.

64. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Rapp BA: **GenBank.** *Nucleic Acids Res* 2002, **30**:17-20.

65. Stoesser G, Baker W, van den Broek A, Camon E, Garcia-Pastor M, Kanz C, Kulikova T, Leinonen R, Lin Q, Lombard V, Lopez R, Redaschi N, Stoehr P, Tuli MA, Tzouvara K, Vaughan R: **The EMBL nucleotide sequence database.** *Nucleic Acids Res* 2002, **30**:21-26.

66. Tateno Y, Imanishi S, Miyazaki S, Fukami-Kobayashi K, Saitou N, Sugawara H, Gojobori T: **DNA Data Bank of Japan (DDBJ) for genome scale research in life sciences.** *Nucleic Acids Res* 2002, **30**:27-30.

67. Wheeler DL, Church DM, Lash AE, Leipe DD, Madden TL: **Database resources of the National Center for Biotechnology Information.** *Nucleic Acids Res* 2001, **29**:11-16.

68. Shumway M, Cochrane G, Sugawara H: **Archiving next generation sequencing data.** *Nucleic Acids Res* 2010, **38**:D870-D871.

69. Cherry JM, Ball C, West S, Juvic G, Schmidt R, Adler C, Dunn B, Dwight S, Riles L, Mortimer RK, Botstein D: **Genetic and physical maps of *Saccharomyces cerevisiae*.** *Nature* 1997, **387**:67-73.

70. Tweedie S, Ashburner M, Falls K, Leyland P, McQuilton P, Marygold S, Millburn G, Osumi-Sutherland D, Schroeder A, Seal R, Zhang H, Consortium FlyBase: **FlyBase: enhancing *Drosophila* Gene Ontology annotations.** *Nucleic Acids Res* 2009, **37**:D555-D559.

71. Bult CJ, Epping JT, Kadin JA, Richardson JE, Blake JA, Mouse Genome Database Group: **The Mouse Genome Database (MGD): mouse biology and model systems.** *Nucleic Acids Res* 2008, **36**:D724-D728.

72. Harris TW, Antoshechkin I, Bieri T, Blasius D, Chan J, Chen WJ, De La Cruz N, Davis P, Duesbury M, Fang R, Fernandes J, Han M, Kishore R, Lee R, Müller HM, Nakamura C, Ozersky P, Petcherski A, Rangarajan A, Rogers A, Schindelman G, Schwarz EM, Tuli MA, Van Auken K, Wang D, Wang X, Williams G, Yook K, Durbin R, Stein LD, Spieth J, Sternberg PW: **WormBase: a comprehensive resource for nematode research.** *Nucleic Acids Res* 2010, **38**:D463-D467.

73. Schwarz EM, Antoshechkin I, Bastiani C, Bieri T, Blasius D, Canaran P, Chan J, Chen N, Chen WJ, Davis P, Fiedler TJ, Girard L, Harris TW, Kenny EE, Kishore R, Lawson D, Lee R, Müller HM, Nakamura C, Ozersky P, Petcherski A, Rogers A, Spooner W, Tuli MA, Van Auken K, Wang D, Durbin R, Spieth J, Stein LD, Sternberg PW: **WormBase: better software, richer content.** *Nucleic Acids Res* 2006, **34**:D475-D478.

74. Jarvie T, Harkins T: **Transcriptome sequencing with the Genome Sequencer FLX system.** *Nat Methods* 2008, **5**:6-8.

75. Pepke S, Wold B, Mortazavi A: **Computational approaches to the analysis of ChIP-seq and RNA-seq data.** *Nat Methods* 2009, **6**:S22-S32.

76. Reinhardt JA, Baltrus DA, Nishimura MT, Jeck WR, Jones CD, Dangl JL: **De novo assembly using low-coverage short read sequence data from the rice pathogen *Pseudomonas syringae* pv. *oryzae*.** *Genome Res* 2009, **19**:294-305.

77. Nagarajan H, Butler JE, Klimes A, Qiu Y, Zengler K, Ward J, Young ND, Methé BA, Palsson BØ, Lovley DR, Barrett CL: **De novo assembly of the complete genome of an enhanced electricity-producing variant of *Geobacter sulfurreducens* using only short reads.** *PLoS One* 2010, **5**:e10922.

78. Tsai IJ, Otto TD, Berriman M: **Improving draft assemblies by iterative mapping and assembly of short reads to eliminate gaps.** *Genome Biol* 2010, **11**:R41.

79. Gregory SG, Barlow KF, McLay KE, Kaul R, Swarbreck D, Dunham A, Scott CE, Howe KL, Woodfine K, Spencer CC, Jones MC, Gillison C, Searle S, Zhou Y, Kokocinski F, McDonald L, Evans R, Phillips K, Atkinson A, Cooper R, Jones C, Hall RE, Andrews TD, Lloyd C, Ainscough R, Almeida JP, Ambrose KD, Anderson F, Andrew RW, Ashwell RI, Aubin K, Babbage AK, Bagguley CL, Bailey J, Beasley H, Bethel G, Bird CP, Bray-AlLEN S, Brown JY, Brown AJ, Buckley D, Burton J, Bye J, Carder C, Chapman JC, Clark SY, Clarke G, Clee C, Cobley V, Collier RE, Corby N, Coville GJ, Davies J, Deadman R, Dunn M, Earthrow M, Ellington AG, Errington H, Frankish A, Frankland J, French L, Garner P, Garnett J, Gay L, Ghori MR, Gibson R, Gilby LM, Gillett W, Glithero RJ, Graham DV, Griffiths C, Griffiths-Jones S, Grocock R, Hammond S, Harrison ES, Hart E, Haugen E, Heath PD, Holmes S, Holt K, Howden PJ, Hunt AR, Hunt SE, Hunter G, Isherwood J, James R, Johnson C, Johnson D, Joy A, Kay M, Kershaw JK, Kibukawa M, Kimberley AM, King A, Knights AJ, Lad H, Laird G, Lawlor S, Leongamornlert DA, Lloyd DM, Loveland J, Lovell J, Lush MJ, Lyne R, Martin S, Mashreghi-Mohammadi M, Matthews L, Matthews NS, McLaren S, Milne S, Mistry S, Moore MJ, Nickerson T, O'Dell CN, Oliver K, Palmeiri A, Palmer SA, Parker A, Patel D, Pearce AV, Peck AI, Pelan S, Phelps K, Phillimore BJ, Plumba R, Rajan J, Raymond C, Rouse G, Saenphimmachak C, Sehra HK, Sheridan E, Showkeen R, Sims S, Skuce CD, Smith M, Steward C, Subramanian S, Sycamore N, Tracey A, Tromans A, Van Helmond Z, Wall M, Wallis JM, White S, Whitehead SL, Wilkinson JE, Willey DL, Williams H, Wilming L, Wray PW, Wu Z, Coulson A, Vaudin M, Sulston JE, Durbin R, Hubbard T, Wooster R, Dunham I, Carter NP, McVean G, Ross MT, Harrow J, Olson MV, Beck S, Rogers J, Bentley DR, Banerjee R, Bryant SP, Burford DC, Burrill WD, Clegg SM, Dhami P, Dovey O, Faulkner LM, Gribble SM, Langford CF, Pandian RD, Porter KM, Prigmore E: **The DNA sequence and biological annotation of human chromosome 1.** *Nature* 2006, **441**:315-321.

80. Coassin S, Brandstatter A, Kronenberg F: **Lost in the space of bioinformatic tools: a constantly updated survival guide for genetic epidemiology. The GenEpi Toolbox.** *Atherosclerosis* 2010, **209**:321-335.

81. Karp PD: **What we do not know about sequence analysis and sequence databases.** *Bioinformatics* 1998, **14**:753-754.

82. Benitez-Paez A: **Considerations to improve functional annotations in biological databases.** *OMICS* 2009, **13**:527-535.

83. Rodriguez-Valle M, Lew-Tabor A, Gondro C, Moolhuijzen P, Vance M, Guerrero FD, Bellgard M, Jorgensen W: **Comparative microarray analysis of *Rhipicephalus (Boophilus) microplus* expression profiles of larvae pre-attachment and feeding adult female stage on *Bos indicus* and *Bos taurus* cattle.** *BMC Genomics* 2010, **11**:437.

84. Huang CQ, Gasser RB, Cantacessi C, Nisbet AJ, Zhong W, Sternberg PW, Loukas A, Mulvenna J, Lin RQ, Chen N, Zhu XQ: **Genomic-bioinformatic analysis of transcripts enriched in the third-stage larva of the parasitic nematode *Ascaris suum*.** *PLoS Negl Trop Dis* 2008, **2**:e246.

85. Geary TG, Thompson DP, Klein RD: **Mechanism-based screening: discovery of the next generation of anthelmintics depends upon more basic research.** *Int J Parasitol* 1999, **29**:105-112.

86. McCarter JP: **Genomic filtering: an approach to discovering novel antiparasitics.** *Trends Parasitol* 2004, **20**:462-468.

87. Krasky A, Rohwer A, Schroeder J, Selzer PM: **A combined bioinformatics and chemoinformatics approach for the development of new antiparasitic drugs.** *Genomics* 2007, **89**:36-43.

88. Caffrey CR, Rohwer A, Ollien F, Marhofer RJ, Braschi S, Oliveira G, McKerrow JH, Selzer PM: **A comparative chemogenomics strategy to predict potential drug targets in the metazoan pathogen, *Schistosoma mansoni*.** *PLoS One* 2009, **4**:e4413.

89. Doyle MA, Gasser RB, Woodcroft BJ, Hall RS, Ralph SA: **Drug target prediction and prioritization: using orthology to predict essentiality in parasite genomes.** *BMC Genomics* 2010, **11**:222.

90. Seib KL, Dougan G, Rappuoli R: **The key role of genomics in modern vaccine and drug design for emerging infectious diseases.** *PLoS Genet* 2009, **5**:e1000612.

91. Woods DJ, Knauer CS: **Discovery of veterinary antiparasitic agents in the 21st century: a view from industry.** *Int J Parasitol* 2010, **40**:1177-1181.

92. Gilleard JS, Woods DJ, Dow JAT: **Model-organism genomics in veterinary parasite drug-discovery.** *Trends Parasitol* 2005, **21**:302-305.

93. Zhong W, Sternberg PW: **Genome-wide prediction of *C. elegans* genetic interactions.** *Science* 2006, **311**:1481-1484.

94. Lee I, Lehner B, Crombie C, Wang W, Fraser AG, Marcotte EM: **A single network comprising the majority of genes accurately predicts the phenotypic effects of gene perturbation in *C. elegans*.** *Nature Genet* 2008, **40**:181-188.

95. Campbell BE, Nagaraj SH, Hu M, Zhong W, Sternberg PW, Ong EK, Loukas A, Ranganathan S, Beveridge I, McInnes RL, Hutchinson GW, Gasser RB: **Gender-enriched transcripts in *Haemonchus contortus*-predicted functions and genetic interactions based on comparative analyses with *Caenorhabditis elegans*.** *Int J Parasitol* 2008, **38**:65-83.

96. Hopkins AL, Groom CR: **The druggable genome.** *Nature* 2002, **1**:727-730.

97. Chang A, Scheer M, Grote A, Schomburg I, Schomburg D: **BRENDA, AMENDA and FRENDA the enzyme information system: new content and tools in 2009.** *Nucleic Acids Res* 2009, **37**:D588-D592.

98. Debrabant A, Joshi MB, Pimenta PF, Dwyer DM: **Generation of *Leishmania donovani* axenic amastigotes: their growth and biological characteristics.** *Int J Parasitol* 2004, **34**:205-217.

99. Lehtinen LE, Birkenheuer AJ, Droleskey RE, Holman PJ: *In vitro* cultivation of a newly recognized *Babesia* sp. in dogs in North Carolina. *Vet Parasitol* 2008, **151**:150-157.

100. Lackovic K, Parisot JP, Sleefs N, Baell JB, Debien L, Watson KG, Curtis JM, Handman E, Street IP, Kedzierski L: Inhibitors of *Leishmania* GDP-mannose pyrophosphorylase identified by high-throughput screening of small-molecule chemical library. *Antimicrob Agents Chemother* 2010, **54**:1712-1719.

101. Chan KY, Matthews KR, Ersfeld K: Functional characterisation and drug target validation of a mitotic kinesin-13 in *Trypanosoma brucei*. *PLoS Pathog* 2010, **6**:pii: e1001050.

102. Suarez CE, McElwain TF: Transfection systems for *Babesia bovis*: a review of methods for the transient and stable expression of exogenous genes. *Vet Parasitol* 2010, **167**:205-215.

103. Bajwa J, Duke SO, Tekwani BL: *Plasmodium falciparum* serine/threonine phosphoprotein phosphatases (PPP): from housekeeper to the 'holy grail'. *Curr Drug Targets* 2008, **9**:997-1012.

104. Price HP, Peltan A, Stark M, Smith DF: The small GTPase ARL2 is required for cytokinesis in *Trypanosoma brucei*. *Mol Biochem Parasitol* 2010, **173**:123-131.

105. Liotta F, Siekierka JJ: Apicomplexa, trypanosoma and parasitic nematode protein kinases as antiparasitic therapeutic targets. *Curr Opin Invest Drugs* 2010, **11**:147-156.

106. Konstantinopoulos PA, Karamouzis MV, Papavassiliou AG: Post-translational modifications and regulation of the Ras superfamily of GTPases as anticancer drugs. *Nature Rev Drug Discov* 2007, **6**:541-555.

107. Lundquist EA: **Small GTPases**. *WormBook* The *C. elegans* Research Community; 2007.

108. van Golen K: **The Rho GTPases in Cancer**. Springer New York Dordrecht Heidelberg; London; 2009.

109. Meena LS: Guanosine triphosphatases as novel therapeutic targets in tuberculosis. *Int J Infect Dis* 2010, **14**:e682-7.

110. Gibbs JB, Oliff A, Kohl NE: Farnesyltransferase inhibitors: Ras research yields a potential cancer therapeutic. *Cell* 1994, **77**:175-178.

111. Gibbs J, Oliff A: The potential of farnesyltransferase inhibitors as cancer chemotherapeutics. *Ann Rev Pharmacol Toxicol* 1997, **37**:143-166.

112. Virtanen SS, Sandholm J, Yegutkin G, Vaananen HK, Harkonen PI: Inhibition of GGTase I and FTase disrupts cytoskeletal organization of human PC-3 prostate cancer cells. *Cell Biol Int* 2010, **34**:815-826.

113. Wieman H, Tøndel K, Anderssen E, Drabøl F: Homology-based modelling of targets for rational drug design. *Mini Rev Med Chem* 2004, **4**:793-804.

114. de Beer TA, Wells GA, Burger PB, Joubert F, Marechal E, Birkholtz L, Louw AJ: Antimalarial drug discovery: *in silico* structural biology and rational drug design. *Infect Disord Drug Targets* 2009, **9**:304-318.

115. Asojo OA, Loukas A, Inan M, Barent R, Huang J, Plantz B, Swanson A, Gouthro M, Meagher MM, Hotez PJ: Crystallization and preliminary X-ray analysis of Na-ASP-1, a multi-domain pathogenesis-related-1 protein from the human hookworm parasite *Necator americanus*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 2005, **61**:391-394.

116. Asojo OA, Goud G, Dhar K, Loukas A, Zhan B, Deumic V, Liu S, Borgstahl GE, Hotez PJ: X-ray structure of Na-ASP-2, a pathogenesis-related-1 protein from the nematode parasite, *Necator americanus*, and a vaccine antigen for human hookworm infection. *J Mol Biol* 2005, **346**:801-814.

117. Lorber DM, Shoichet BK: Hierarchical docking of databases of multiple ligand conformations. *Curr Top Med Chem* 2005, **5**:739-749.

118. McInnes C, Fischer PM: Strategies for the design of potent and selective kinase inhibitors. *Curr Pharm Des* 2005, **11**:1845-1863.

119. Villoutreix BO, Renault N, Lagorce D, Sperandio O, Montes M, Miteva MA: Free resources to assist structure-based virtual ligand screening experiments. *Curr Protein Pept Sci* 2007, **8**:381-411.

120. Yarnitzky T, Levit A, Niv MY: Homology modeling of G-protein-coupled receptors with X-ray structures on the rise. *Curr Opin Drug Discov Devel* 2010, **13**:317-325.

121. Cohen P: Protein kinases - the major drug targets of the twenty-first century? *Nat Rev Drug Discov* 2002, **1**:309-315.

122. Dissous C, Ahier A, Khayath N: Protein tyrosine kinases as new potential targets against human schistosomiasis. *Bioessays* 2007, **29**:1281-1288.

123. Rastelli G, Paccioni S, Sirawaraporn W, Sirawaraporn R, Parenti MD, Ferrari AM: Docking and database screening reveal new classes of *Plasmodium falciparum* dihydrofolate reductase inhibitors. *J Med Chem* 2003, **46**:2834-2845.

124. Peng H, Huang N, Qi J, Xie P, Xu C, Wang J, Yang C: Identification of novel inhibitors of BCR-ABL tyrosine kinase via virtual screening. *Bioorg Med Chem Lett* 2003, **13**:3693-3699.

125. Vangrevelinghe E, Zimmermann K, Schoepfer J, Portmann R, Fabbro D, Furet P: Discovery of a potent and selective protein kinase CK2 inhibitor by high-throughput docking. *J Med Chem* 2003, **46**:2656-2662.

126. Hammami R, Fliss I: Current trends in antimicrobial agent research: chemo and bioinformatics approaches. *Drug Discov Today* 2010, **15**:540-546.

127. Dorato MA, Buckley LA: Toxicology testing in drug discovery and development. *Curr Prot Toxicol* 2007, **31**:19.1.1-35.

128. Allocco JJ, Donald R, Zhong T, Lee A, Tang YS, Hendrickson RC, Liberator P, Nare B: Inhibitors of caseine kinase 1 block the growth of *Leishmania major* promastigotes *in vitro*. *Int J Parasitol* 2006, **36**:1249-1259.

129. Knobloch J, Kunz W, Greveling CG: Herbimycin A suppresses mitotic activity and egg production of female *Schistosoma mansoni*. *Int J Parasitol* 2006, **36**:1261-1272.

130. World Health Organization: **World Health Organization Report. Changing History**. Geneva; Switzerland 2004.

131. Barbieri CL: Immunology of canine leishmaniasis. *Parasite Immunol* 2006, **28**:329-337.

132. Belkaid Y, Sun CM, Bouladoux N: Parasites and immunoregulatory T cells. *Curr Opin Immunol* 2006, **18**:406-412.

133. Baneth G, Koutinas A, Solano-Gallego L, Bourdeau P, Ferrer L: Canine leishmaniosis - new concepts and insights on an expanding zoonosis: part one. *Trends Parasitol* 2008, **24**:324-330.

134. Woods D, Butler C, Williams T, Greenwood K: Receptor-based discovery strategies for insecticides and parasiticides: a review. *Adv Exp Med Biol* 2010, **692**:1-9.

135. Smout MJ, Kotze AC, McCarthy JS, Loukas A: A novel high throughput assay for antihelminthic drug screening and resistance diagnosis by real-time monitoring of parasite motility. *PLoS Negl Trop Dis* 2010, **4**:e885.

136. Engel JC, Ang KK, Chen S, Arkin MR, McKerrow JH, Doyle PS: Image-based high-throughput drug screening targeting the intracellular stage of *Trypanosoma cruzi*, the agent of Chagas' disease. *Antimicrob Agents Chemother* 2010, **54**:3326-3334.

137. Siqueira-Neto JL, Song OR, Oh H, Sohn JH, Yang G, Nam J, Jang J, Cechetto J, Lee CB, Moon S, Genovesio A, Chatelain E, Christophe T, Freitas-Junior LH: Antileishmanial high-throughput drug screening reveals drug candidates with new scaffolds. *PLoS Negl Trop Dis* 2010, **4**:e675.

doi:10.1186/1756-3305-4-53

**Cite this article as:** Gasser et al.: Major prospects for exploring canine vector borne diseases and novel intervention methods using 'omic technologies. *Parasites & Vectors* 2011 **4**:53.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

